1. Quality audit records are kept on numbers of major and minor failures of the construction process in a big firm. For a process of this type, two random variables
X = the number of major failures and
Y = the number of minor failures
can be described at least approximately by the accompanying joint distribution.
(2 points each)

y \ x	0	1	2	Total
0	0.15	0.05	0.01	0.21
1	0.10	0.10	0.04	0.24
2	0.10	0.14	0.04	0.28
3	0.10	0.13	0.04	0.27
Total	0.45	0.42	0.13	1

a. Is the table above a valid joint pmf for X and Y? Why
b. Find the marginal probability mass functions for X and Y.
c. Are X and Y independent?
d. Find the expected value and the variance of X, E(X) and Var(X)
e. Find the expected value and the variance of Y.
f. Find the Cov(X,Y) and Corr(X,Y). Hint: Cov(X,Y) = E(XY) – E(X) ·E(Y), Corr(X,Y) = Cov(X, Y)/[$\sigma_x \cdot \sigma_y$]
g. Find the conditional probability function for Y given that X=0 that is there are no major failures.
h. What is the expected number of minor failures given that there were no major failures?
i. Suppose that demerits are assigned to devices of this type according to the formula D=2X+Y. Find the marginal probability mass function for D.

2. For the question below, the *italicized* statement is either True or False. Suppose y ~ N(μ; σ^2) and you perform a test of $H_0 : \mu = 10$ vs. $H_A : \mu \neq 10$ with a random sample of size 9. You observe $\bar{y}$ and s^2 and the subsequent t-test results in a p-value of 0.04. Suppose you now test the same hypotheses with a random sample of size 36. Suppose you observe the same values for $\bar{y}$ and s^2 as before. *You will certainly reject the null hypothesis at $\alpha = 0.05$.* Indicate whether the statement is True or False and provide a justification for your response. (10 points)

3. Several numbers are missing in the attached ANOVA table. The data were obtained in a regression experiment with 16 observations. The regression line is expressed as follows: $y = \beta_0+\beta_1 x$.
   a Please fill in the seven missing numbers. (14 points)
   b What does the statement "Prob. > F, 0.001" mean? (Hint: Start with the null hypothesis of the regression equation and think through this question) (3 points)

Source	SS	DF	MS	F	Prob. > F
Model	2059.78	(1)	(3)	(6)	0.0001
Error	974.65	(2)	(5)		
Total	3034.43	(2)			

R^2 = (7)

4. A civil engineer conducted a study comparing the effect of various additives on the engineering property of a material. A total of 5 additives were used. For each additive, 6 samples were randomly selected for testing. The engineering property was determined after an additive was mixed with the material. The following data are available. The sample variances for the varieties are s_1^2 = 0.258, s_2^2 = 0.755; s_3^2 = 0.346; s_4^2 = 0.608 and s_5^2 = 0.460. Also, SSTr = 17.093.

   a. You are asked to complete the ANOVA table. (10 points)
   b. Please determine the table F value. Then, you have to carry out a test of the null hypothesis that the population means of all 5 varieties are equal. Interpret your results. (5 points)

5. The worker of a construction company has the mean of 74 working hours and the standard deviation of 5 hours. What is the probability that the total hours of (a random sample of) 36 workers is less than 2,700 hours? (10 points)

6. The tensile strength of a material is modeled by a normal distribution with mean 3.84 MPa and standard deviation 0.85 MPa.
   a The materials with the highest 15% of tensile strength are considered to have unusually high levels for stress-resistant materials. What threshold value of strength must a material exceed in order to be considered to have unusually high strength levels for a stress-resistant material? (10 points)
   b Forty-five samples are selected at random and tested for their tensile strength. What is the probability that the average strength for the 45 samples is greater than 4.25 MPa? (10 points)

7. A construction company is thinking about putting in a bid for four projects. The first two projects have expected returns of μ_1 = 5% per year with a standard deviation of 1% whereas the last two projects have expected returns of μ_2 = 10% per year with a standard deviation of 2%. The company plans to split their investment equally among the four projects so that the parameter of interest is the true mean return

$$\theta = 0.25\mu_2 + 0.25\mu_2 + 0.25\mu_3 + 0.25\mu_4 = 0.5(\mu_1 + \mu_2);$$

which would be 7.5%. If the yearly returns of the 4 projects are denoted by X_1, X_2, X_3, X_4, the company considers two competing estimators for the mean yearly return.

$$\theta_1 = 0.25X_1 + 0.25X_2 + 0.25X_3 + 0.25X_4$$

and

$$\theta_2 = 0.4X_1 + 0.4X_2 + 0.1X_3 + 0.1X_4$$

Evaluate the biases of the two estimators and give their variances and mean squared errors. Which estimator should the company prefer and why? Hint: MSE $(\hat{\theta})$ = Var $(\hat{\theta})$ + [bias $(\hat{\theta})]^2$ (10 points)

Values Provided for Your Calculations

Student t Distribution

DF	Upper-Tail Area α 0.05	0.025
8	1.860	2.306
9	1.833	2.262

Normal Distribution

z	1.04	1.2	1.64	2.33	2.58	3.24
$\Phi(z)$	0.85	0.88	0.95	0.99	0.995	0.999

F Distribution (α = 0.01)

		ν_1 = numerator df 4	5	6
ν_2 = denominator df	24	6.59	5.98	5.55
	25	6.49	5.89	5.46
	26	6.41	5.80	5.38
	27	6.33	5.73	5.31
	28	6.25	5.66	5.24
	29	6.19	5.59	5.18
	30	6.12	5.53	5.12